

Online Research @ Cardiff

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository: <https://orca.cardiff.ac.uk/id/eprint/89013/>

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Pickup, David ORCID: <https://orcid.org/0000-0003-2894-0822>, Sun, Xianfang ORCID: <https://orcid.org/0000-0002-6114-0766>, Rosin, Paul L. ORCID: <https://orcid.org/0000-0002-4965-3884> and Martin, Ralph Robert 2016. Skeleton-based canonical forms for non-rigid 3D shape retrieval. Computational Visual Media 2 (3) , pp. 231-243. 10.1007/s41095-016-0045-5 file

Publishers page: <http://dx.doi.org/10.1007/s41095-016-0045-5>
<<http://dx.doi.org/10.1007/s41095-016-0045-5>>

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies.

See

<http://orca.cf.ac.uk/policies.html> for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.



Skeleton-based canonical forms for non-rigid 3D shape retrieval

David Pickup¹✉, Xianfang Sun¹, Paul L. Rosin¹, and Ralph R. Martin¹

© The Author(s) 2015. This article is published with open access at Springerlink.com

Abstract The retrieval of non-rigid 3D shapes is an important task. A common technique is to simplify this problem to a rigid shape retrieval task by producing a bending invariant canonical form for each shape in the dataset to be searched. It is common for these techniques to attempt to “unbend” a shape by applying multidimensional scaling to the distances between points on the mesh, but this leads to unwanted local shape distortions. We instead perform the unbending on the skeleton of the mesh, and use this to drive the deformation of the mesh itself. This leads to a computational speed-up and less distortions of the local details of the shape. We compare our method against other canonical forms and our experiments show that our method achieves state-of-the-art retrieval accuracy in a recent canonical forms benchmark, and only a small drop in retrieval accuracy over state-of-the-art in a second recent benchmark, while being significantly faster.

Keywords canonical forms, shape retrieval, skeletons, pose invariance.

1 Introduction

The task of example based retrieval of non-rigid objects is both a key problem to solve and a challenging one. There are increasing numbers of 3D shape collections being created, so the ability to search these collections is an increasingly important task. There have been many successes in the retrieval of rigid objects, with methods such as view based techniques proving very successful [14]. The problem is that many of these techniques cannot be applied to non-rigid shape retrieval. To address this issue Elad and

Kimmel [9] proposed a bending invariant 3D embedding of a mesh, named a *canonical form*. The canonical form of a mesh effectively standardises its pose, and therefore when a canonical form is computed for each shape in a dataset the non-rigid retrieval problem becomes a rigid retrieval problem. This means that any of the wide range of rigid shape retrieval methods available are then able to perform retrieval on this data. There are two issues with the canonical form method by Elad and Kimmel. The first is that it requires the geodesic distance between all pairs of vertices to be computed, which has a *super-quadratic* computational complexity. The second issue is that the small scale local details of the shapes are lost in the canonical form.

In this paper we address both these issues by applying the method by Elad and Kimmel to the skeleton of the mesh, rather than to the mesh itself. The pose of the mesh is then deformed by the pose of the resulting canonical skeleton. This is far more efficient because the skeleton of a mesh contains far fewer vertices than the mesh itself, resulting in far fewer geodesic distance computations. Less shape details are distorted, because the method effectively produces a set of canonical angles at the articulated joints of the shape, and the shape deformations are localised to these joint regions. This leads to an increased retrieval performance on a recent canonical forms benchmark [21].

The structure of our paper is as follows. Section 2 outlines the related work in this area, Section 3 describes the technical details of our method, Section 4 presents the results of our experiments, and we make our conclusions in Section 5.

2 Related Work

There are many works which aim to solve the retrieval problem for rigid shapes, such as *lightfield* descriptors [7] and *spin images* [12]. We refer readers to [14] and [28] for detailed reviews of this field of research.

In recent years more research has concentrated on the problem of retrieving non-rigid models. Several methods extract local features from a mesh to compute a

¹ School of Computer Science and Informatics, Cardiff University, Cardiff, CF24 3AA, UK. E-mail: PickupD@Cardiff.ac.uk, SunX2@Cardiff.ac.uk, RosinPL@Cardiff.ac.uk, MartinRR@Cardiff.ac.uk.

Manuscript received: xxxx-xx-xx; accepted: xxxx-xx-xx.

shape descriptor [6], including *meshSIFT* [26], *conformal factors* [3], *area projection transforms* [10] and *heat kernel signatures* [27]. Computing histograms of local features has proven to perform very well in recent retrieval benchmarks [18, 20].

Graphs have also been used as shape descriptors. Hilga *et al.* [11] used *multiresolution Reeb graphs* to match the topology between 3D shapes, and Sfikas *et al.* [23] proposed formulating a graph based on conformal factors [3].

Matching global information of 3D shapes has also proved successful. Reuter *et al.* [22] demonstrated that the Laplace-Beltrami spectra can be used as a shape descriptor, which they name *ShapeDNA*. Various global descriptors can also be extracted from the geodesic distance matrix of a mesh, as shown by Smeets *et al.* [25].

For a more detailed review of the latest non-rigid retrieval methods, we refer the reader to the recent SHREC benchmarks [18, 20].

The use of canonical forms to normalise the pose of non-rigid shapes was first proposed by Elad and Kimmel [9]. They use *multidimensional scaling* to map the geodesic distances of a mesh into 3D Euclidean distances. Several variations to this method have been proposed. Shamai *et al.* [24] accelerate the classical MDS procedure using their proposed *Nyström Multidimensional Scaling* framework. Lian *et al.* [17] attempt to preserve the features of a mesh by segmenting the original mesh and transforming each segment to its location in the canonical mesh computed using Elad and Kimmel's method, thus correcting some of the local shape distortions. Wang and Zha [29] speed up the canonical form computation by only computing geodesic distances between all pairs of a set of detected feature points, and unbending the mesh by creating target axes used to align sets of geodesic contours. Pickup *et al.* [19] also use feature points, but restrict their number to the square root of the number of mesh vertices. They maximise the distance between pairs of these feature points whilst preserving the mesh's edge lengths. Boscaini *et al.* [5] proposed a method which assigns a repelling electrical charge at each vertex of the mesh to form a canonical form. They are also able to correct certain very small localised topological errors in the mesh, by cutting parts of the mesh which are likely to have been incorrectly joined. Their method is faster than Elad and Kimmel's, but still suffers from distorting local shape details. There is also work on parallelising or speeding up the computation of geodesic distances [8, 31]. Recent benchmarks [15, 18] have shown that using canonical forms along with the view-based shape retrieval method by Lian *et al.* [16] performs very competitively compared with other non-rigid retrieval approaches.

Our method differs from those above by normalising

the pose of the shape's skeleton, rather than performing the computation on the mesh vertices. This causes less distortion to the local shape details compared with other methods, whilst providing a practical level of efficiency.

3 Method

To describe the workings of our method, we first give an overview of the canonical form work by Elad and Kimmel [9] in Section 3.1 as our work builds from this approach. We then detail our novel skeleton-based approach in Section 3.2.

3.1 Background

Our method extends the canonical form work by Elad and Kimmel [9]. Their method transforms the mesh so that the geodesic distance between all pairs of vertices are mapped to Euclidean distances. To accomplish this they first compute the geodesic distance between all pairs of mesh vertices using the *fast marching method* [13]. Next they use multidimensional scaling to calculate a new set of vertex positions, where the Euclidean distance between each pair of vertices is as close as possible to the already computed geodesic distances. They show results with three different multidimensional scaling techniques, but the one which tends to provide the best results, and which we use in our work, solves the multidimensional scaling problem by minimising the following least squares functional:

$$S(X) = \sum_{i=1}^{N-1} \sum_{j=i+1}^N w_{i,j} (\delta_{i,j} - d_{i,j}(X))^2, \quad (1)$$

where N is the number of vertices, $w_{i,j}$ are weighting coefficients, $\delta_{i,j}$ is the geodesic distance between vertices i and j of the original mesh, and $d_{i,j}$ is the Euclidean distance between vertices i and j of the resulting canonical mesh X . This functional is minimised using the *SMACOF* (scaling by maximising a convex function) algorithm [4].

This method is computationally expensive to compute, as the geodesic distances are calculated in $O(N^2 \log N)$ time and each iteration of the SMACOF algorithm in $O(N^2)$ time. As this method works on the vertices of the mesh, it deforms the details of its shape as well as normalising its pose. These details may be important when analysing the shape of the object.

3.2 Skeleton Canonical Forms

The purpose of computing a canonical form is to normalise the pose of a 3D object. The pose of an object can be defined as the articulation of the object's skeleton. This leads us to our method, which normalises the pose of an object by transforming it so that its skeleton is in a normalised pose. The stages of our method are depicted in Figure 1. We first extract the skeleton of the mesh

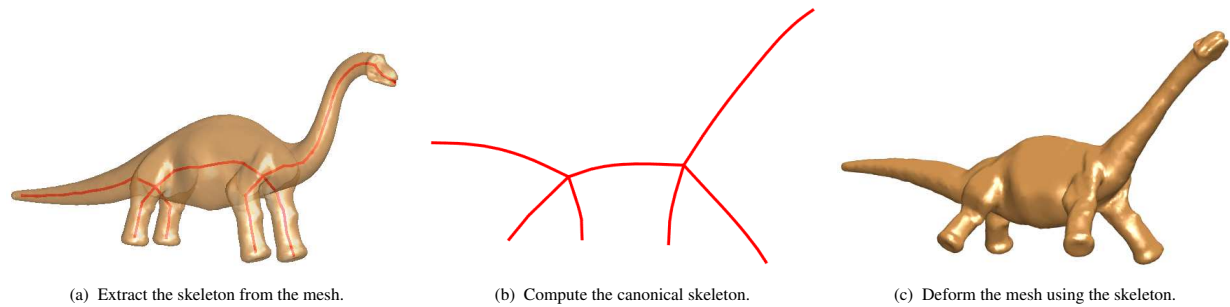


Fig. 1 Outline of our method.

(Section 3.2.1), then transform the skeleton into a canonical form (Section 3.2.2), and finally deform the mesh according to the skeleton transformation (Section 3.2.3).

3.2.1 Skeleton Extraction

We first extract the mesh's skeleton using the method by Au *et al.* [1]. This method works by first contracting the mesh to a zero-volume skeletal shape using Laplacian smoothing. The mesh is then converted to a 1D curve skeleton by removing all the mesh faces while preserving the skeletal topology.

The skeleton is then refined by merging junctions with neighbouring joints, if the merged junction has a better "centeredness". A junction is defined as a joint attached to three or more bones. The centeredness of a junction is defined as the standard deviation of the distances between the junction's position and the position of each vertex assigned to that junction during the skeleton extraction process. A junction is merged with a neighbour if $\sigma' < 0.9\sigma$, where σ' and σ are the centeredness of the merged and original junctions respectively. Finally the skeleton joints are repositioned to better centre them in the mesh.

Please see the original paper for the full details. An example of the resulting skeleton is shown in Figure 1(a).

3.2.2 Skeleton Transformation

Next we apply the canonical form method by Elad and Kimmel [9] to the skeleton. We use Dijkstra's algorithm to compute the geodesic distances between all the joints of the skeleton, and then use Equation 1 for the multidimensional scaling. An example of this is shown in Figure 1(b). This still has a high time complexity, but it is related to the number of joints in the skeleton instead of the number of mesh vertices. In practice this number is significantly smaller, and therefore computing the canonical form of a skeleton has a significantly shorter runtime.

3.2.3 Mesh Deformation

Finally we deform the mesh according to the canonical transformation of its skeleton using the method by Yan *et al.* [30]. We use this method as it is simple to implement,

fully automatic, does not require any vertex weights to be assigned, and does not have any parameters which require training or manual-tuning. This works by first assigning each triangle to a bone of the original skeleton, and then calculating the transformation of each bone between the original and canonical skeletons. A sparse linear system is then solved, which transforms each triangle according to the transformation of its assigned bone whilst preserving the mesh's connectivity. For the full details of this method, please refer to the original paper.

Yan *et al.*'s method takes into account the translation, rotation and scaling of the skeleton when deforming the mesh, but we ignore scaling and translation as we only care about transforming the articulation of the mesh, and we do not want any stretching of the skeleton caused by the canonicalisation to be transferred to the mesh. We also use our own method for assigning triangles to bones, because the method by Yan *et al.* requires that the ends of the skeleton protrude outside the mesh, which is not true for our skeletons, and because we can retain information from the skeleton extraction procedure to make this assignment.

The skeleton extraction method results in each vertex being assigned to a joint of the skeleton [1]. This assignment is based on which joint each vertex was collapsed into during the skeleton extraction method. We use these assignments to assign each triangle to a bone of the skeleton, as required by Yan *et al.*'s deformation method. For each joint we find all the triangles which have at least one vertex assigned to that joint. Each triangle is then assigned to one of the bones connected to that joint. To determine which of these bones a particular triangle is assigned to, we first calculate a plane which bisects each pair of bones that meet at the joint. A triangle is assigned to a particular bone if two or more of its vertices lie on that bone's side of all the bisection planes between it and the other bones. This is illustrated in Figure 2. If a triangle does not meet the assignment criteria for any of the bones, then one of its neighbours is randomly selected and it is given the same bone assignment as that neighbour. This is outlined in

Algorithm 1.

Algorithm 1 Algorithm for assigning triangles to bones.

Input: mesh, skeleton, assignment of vertices to joints.

Output: each triangle is assigned to a bone.

```

for all joints  $j$  in skeleton do
  bisectionPlanes  $\leftarrow \emptyset$ 
  for all bones  $b_1$  connected to  $j$  do
    for all bones  $b_2$  connected to  $j$ , where  $b_1 \neq b_2$  do
      bisectionPlanes  $\leftarrow$  plane bisecting  $b_1$  and  $b_2$ 
    end for
  end for
  for all triangles  $t$  with a vertex assigned to  $j$  do
    for all bones  $b$  connected to  $j$  do
      if  $\geq 2$  vertices of  $t$  fall on  $b$ 's side of all bisectionPlanes then
        assign  $t$  to  $b$ 
      end if
    end for
    if  $t$  is unassigned then
      randomly select a neighbouring triangle  $n$  of  $t$ 
      copy bone assignment of  $n$  to  $t$ 
    end if
  end for
end for

```

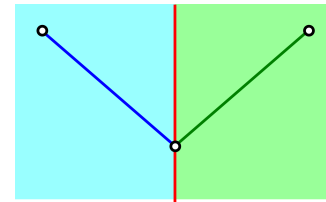
An example of a resulting canonical mesh is shown in Figure 1(c). As you can see, the mesh has been placed into the canonical pose of the skeleton, but with very little shape distortion, hence preserving the surface details.

4 Experiments

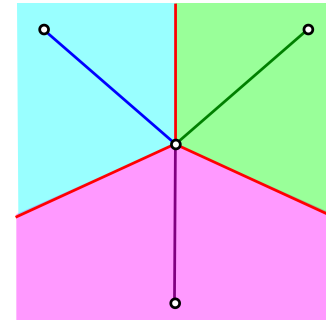
We compare our skeleton canonical forms to several other methods using the two most recently developed publicly available datasets for benchmarking. Firstly we present our retrieval performance on the SHREC'15 canonical forms benchmark [21] in Section 4.1. Secondly we present the results of our experiments on the SHREC'15 non-rigid retrieval benchmark [18] in Section 4.2. Finally we show some limitations of our method in Section 4.3.

4.1 SHREC'15 Canonical Forms Benchmark

Our method was recently entered into the SHREC'15 canonical forms benchmark [21]. The purpose of this benchmark was to compare the effectiveness of different methods at producing canonical forms for 3D shape retrieval. The dataset used for this benchmark contains 100 meshes, split into 10 different shape classes. Each shape class contains a mesh in 10 different non-rigid poses. The average number of vertices per mesh is 21,141. The dataset contains models from both the SHREC'11 non-rigid benchmark [15], which provides a wide range of shape classes, and the SHREC'14 non-rigid humans



(a) Two neighbour example.



(b) Three neighbour example.

Fig. 2 Two dimensional illustration of the assignment regions for bone assignment. The separation planes are shown in red, and the assignment regions are illustrated in the same colour as the corresponding bone. If at least two vertices of a triangle fall within an assignment region, the triangle is assigned to the corresponding bone.

Tab. 1 Comparison of methods on the SHREC'15 canonical forms benchmark [21] using a view-based retrieval method [16]. Original meshes refers to performing retrieval without using canonical forms. Our method achieves the highest retrieval score on three of the four measures.

	NN	FT	ST	DCG
Original Meshes	0.50	0.567	0.702	0.753
Classic MDS	0.73	0.597	0.741	0.796
Fast MDS	0.66	0.590	0.718	0.789
Least Squares MDS	0.75	0.694	0.829	0.838
Non-Metric MDS	0.77	0.687	0.811	0.836
GPS	0.72	0.556	0.697	0.783
Euclidean Random	0.54	0.640	0.783	0.793
Euclidean Normalised	0.61	0.673	0.796	0.816
Least Squares MDS B	0.66	0.662	0.788	0.813
Our method - simplified meshes	0.74	0.682	0.791	0.825
Our method - full meshes	0.77	0.714	0.824	0.849

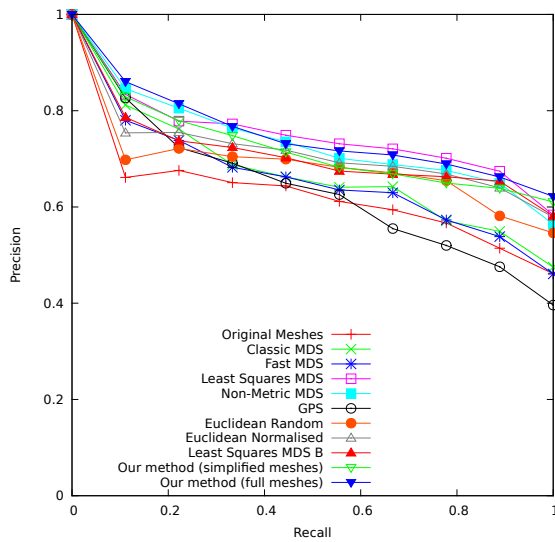


Fig. 3 Precision-recall curves for each method tested on the SHREC'15 canonical forms benchmark [21] using a view-based retrieval method [16]. Our method achieves the best precision for low and high recall values, falling below least-squares MDS for mid-range recall values.

benchmark [20], where the details of the shapes are important for distinguishing between them. Our method was one out of ten methods which took part. All the canonical forms from each method were input to a view-based retrieval method [16] to test their effectiveness for non-rigid retrieval. Here we update our results for this benchmark, as we have rewritten some of our code to improve the speed of our implementation and therefore no longer require the largest meshes (~60,000 vertices) contained in the dataset to be simplified, which was done for the original benchmark paper. This speed-up is achieved purely through changes to our code, with no alterations to the actual algorithms used. The MDS based methods tested all use simplified meshes with 2,000 vertices as input, as these methods are too computationally expensive to compute the canonical forms of the full resolution meshes in a reasonable amount of time. All the other methods use the full resolution meshes as input. The retrieval results are shown in Table 1. All the performance measures produce a value in the interval $[0, 1]$ and are defined as:

Nearest Neighbour (NN) The fraction of closest matches which are members of the query model's class.

First Tier (FT) The fraction of models which are members of the query model's class that appear within the top C closest matches, where C is the number of models in the query's class.

Second Tier (ST) The fraction of models which are members of the query model's class that appear within

the top $2C$ closest matches, where C is the number of models in the query's class.

Discounted Cumulative Gain (DCG) This weights correct matches more if they are higher in the list of retrieved results. The DCG is computed by first assigning each model G_i in the retrieved list G a value of 1 if it is a member of the query's class, and 0 if it is not. An initial DCG is then computed as

$$DCG_i = \begin{cases} G_i & i = 1, \\ DCG_{i-1} + G_i / \log_2(i) & \text{otherwise.} \end{cases} \quad (2)$$

The final result where $i = N$, where N is the number of models in the dataset, is divided by the optimal DCG

$$DCG = \frac{DCG_N}{1 + \sum_{i=2}^C 1 / \log_2(i)}, \quad (3)$$

where C is the number of models in the query's class.

The original submission of our method gave a very competitive performance, achieving the third highest retrieval scores behind the least-squares and non-metric MDS methods which use the mesh's geodesic distances. Our updated result, which uses the full resolution meshes, raises our ranking to achieve state-of-the-art results with the highest retrieval performance on three of the four performance measures. This performance increase is likely due to the full resolution meshes containing important details which are lost when the mesh is simplified. This highlights the importance of being able to efficiently handle meshes with a large number of vertices, and using a canonical form method which preserves the local details. We also show the precision-recall curve [2] of each method in Figure 3. Our method achieves the best precision for low and high recall values, falling below least-squares MDS for mid-range recall values as reflected by the slightly lower second tier measure in Table 1.

4.2 SHREC'15 Non-Rigid Shape Retrieval Benchmark

The SHREC'15 non-rigid benchmark [18] contains a far larger number of 3D models, and therefore we also perform experiments on this dataset. The purpose of this benchmark was to compare the state-of-the-art methods in non-rigid shape retrieval. This dataset contains 1,200 meshes, split into 50 different shape classes. Each shape class contains a mesh within 24 different non-rigid poses. Four meshes in each shape class contain topological errors, such as disconnected components, or unwanted connections (Figure 9). The average number of vertices per mesh is 9,607. We compare our canonical forms against those submitted to the SHREC'15 canonical forms

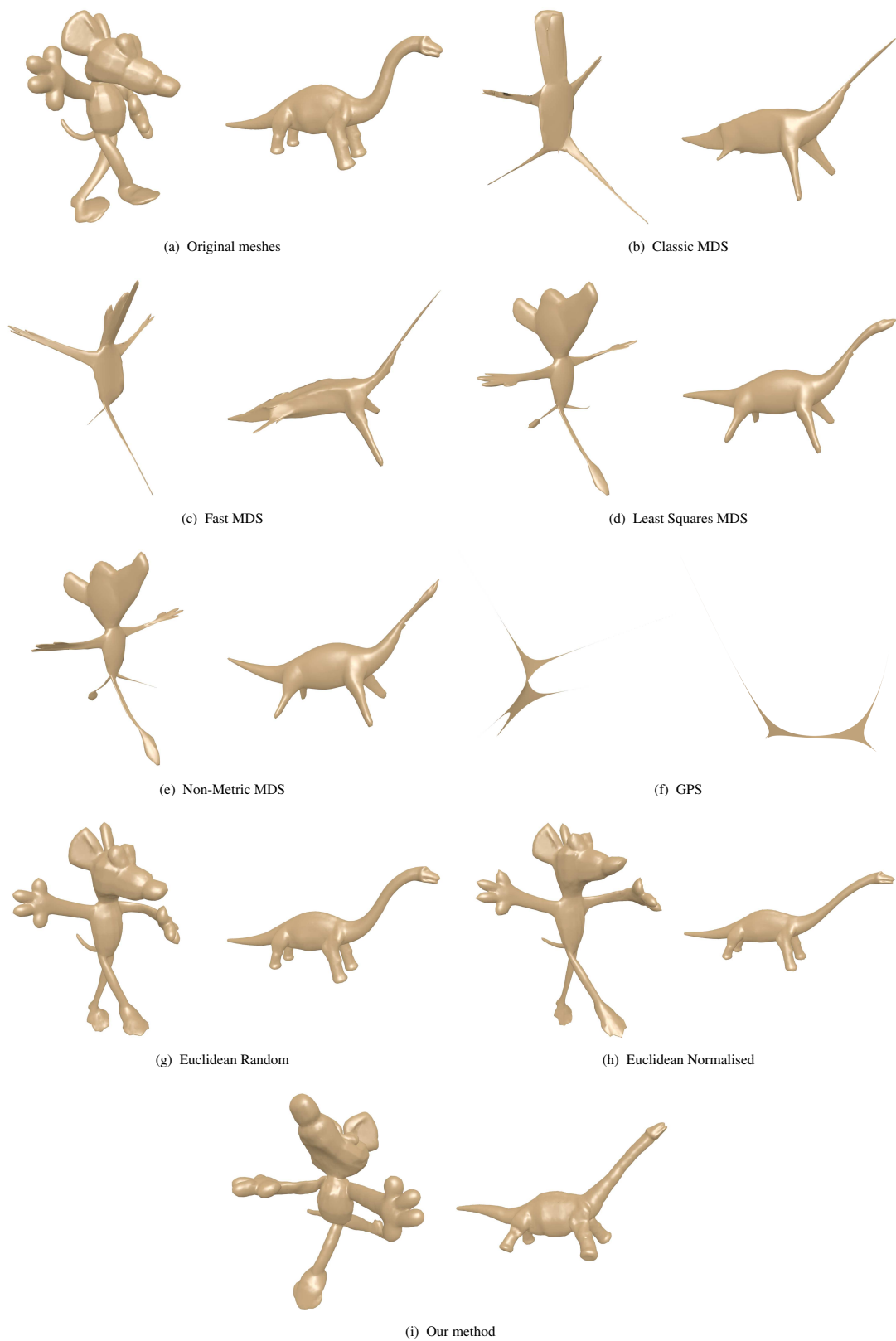


Fig. 4 Example canonical forms for each method tested on the SHREC'15 non-rigid dataset [18].

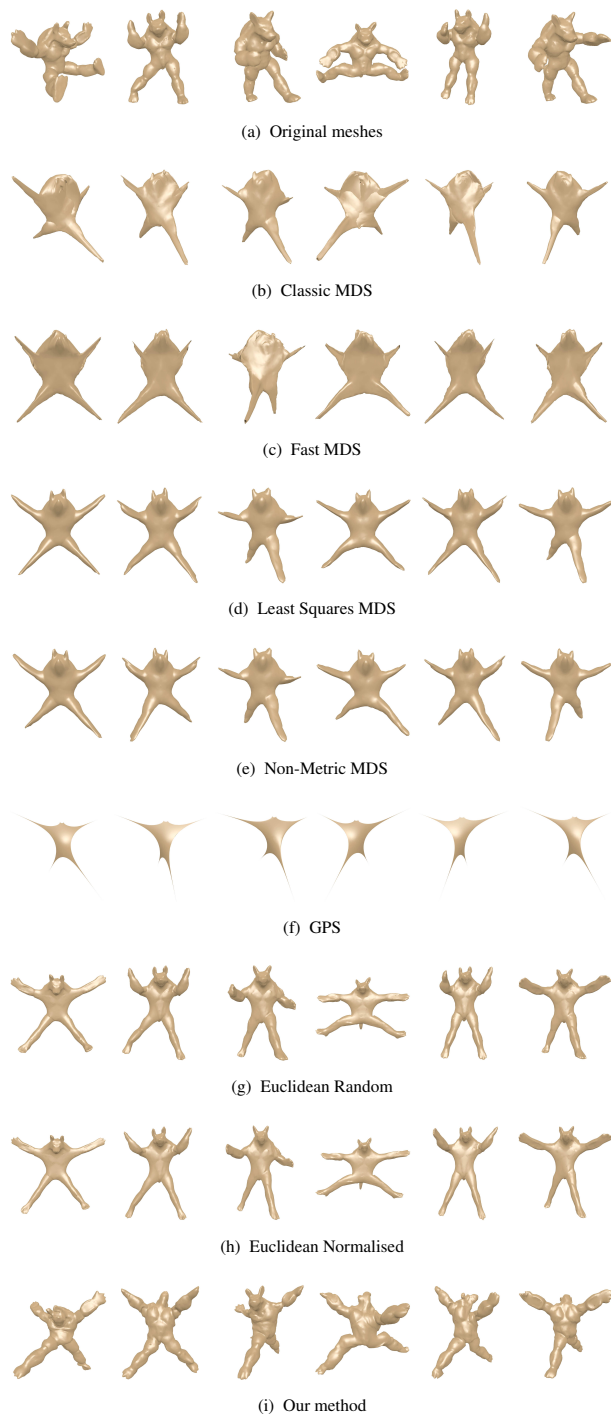


Fig. 5 Example of six canonical forms of the same mesh for each method tested on the SHREC'15 non-rigid dataset [18].

benchmark [21], as we have their implementations. We leave out the least squares MDS B method, as it is simply the same as the least squares method, but with an early termination from the MDS algorithm which has shown to decrease the quality of the canonical forms [21].

Figure 4 shows two example meshes and their associated canonical forms produced using each method. It is noticeable that the GPS method and the MDS based methods severely distort the local shape details. The Euclidean canonical form methods cause slightly less shape distortions, but sometimes fail to completely stretch out the limbs of the shape. Our skeleton method achieves a similar pose to the MDS based methods, but with the least shape distortions of all the methods.

Figure 5 shows the canonical forms for six meshes from the “armadillo” shape class of the SHREC'15 non-rigid dataset, for each of the methods. Our method produces a similarly consistent pose to least squares MDS, but with less shape distortions. Our method however does exhibit some extra inconsistency with respect to the pose of the head of the armadillo. The classic MDS, Fast MDS, and GPS methods distort the shape so much it becomes almost unrecognisable. The Euclidean-based methods distort the shape details less than all methods except ours, but do not produce as consistent a canonical pose.

Some of the meshes contained within this dataset have topological errors to increase the difficulty of the retrieval challenge. One type of error present in some objects is that the mesh is disconnected into two or three different components. The Euclidean distance based methods by Pickup *et al.* [19] do not require any modification for this, but all the other methods fail on these meshes. For the MDS and GPS methods, we therefore delete all but the largest component. For our skeleton method, we test two different solutions. The first is to join the separate skeletons for each component by merging the closest joints between components, and the second method is identical to the solution for the MDS and GPS methods. Figure 6 shows an example of this problem, where the mouse's head is disconnected from its body. The methods which only keep the largest component therefore produce canonical forms without the presence of the mouse's head. The Euclidean distance based methods separate out the head and the body, as there are no edge connections keeping them together. Our method which connects the skeletons, places the head in an odd position. This is because, although the skeletons are connected, the method has no mesh connections to preserve between the head and the body.

All the canonical forms from each method were input to a view-based retrieval method [16] to test their effectiveness for non-rigid retrieval. The retrieval results are shown

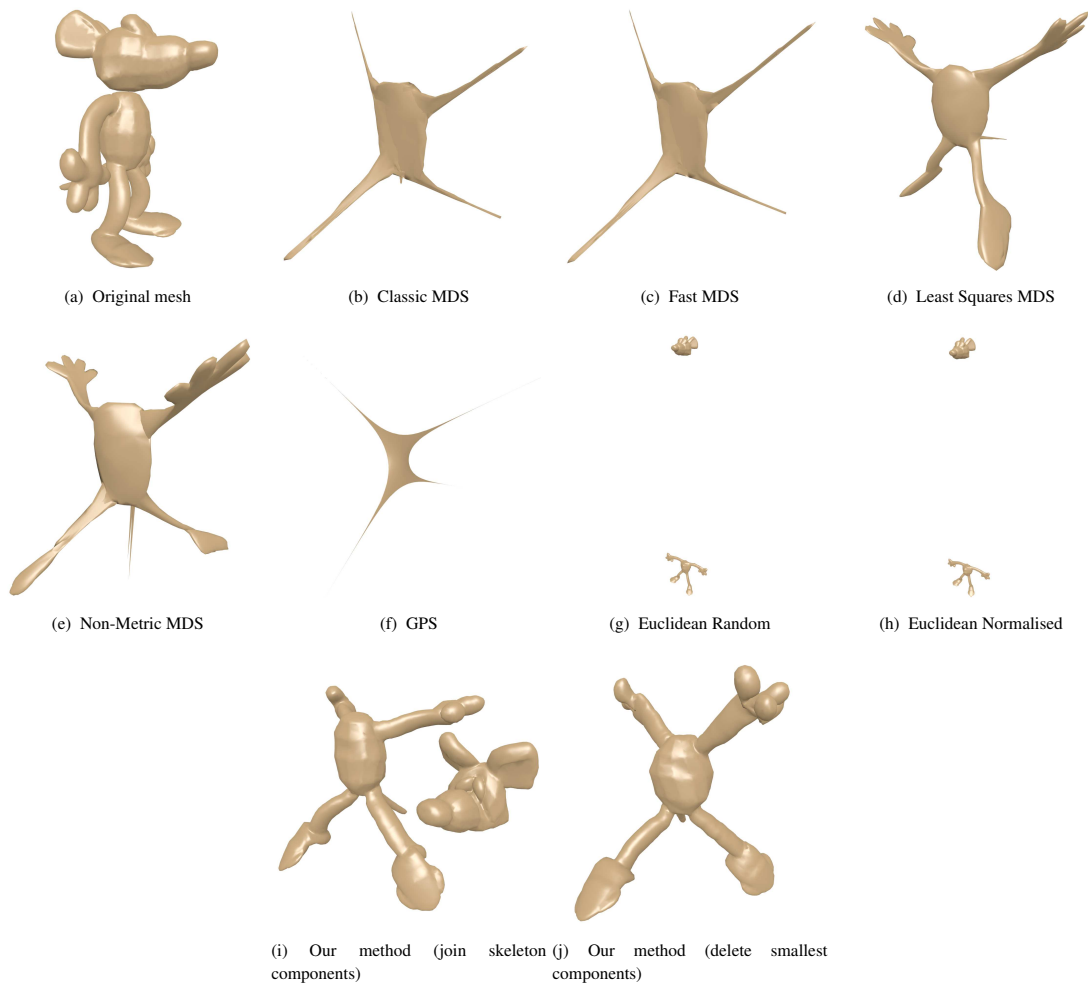


Fig. 6 Canonical forms for a mesh with disconnected components.

Tab. 2 Comparison of methods on the SHREC'15 non-rigid benchmark [18] using a view-based retrieval method [16]. Original meshes refers to performing retrieval without using canonical forms. Our method achieves the third highest retrieval performance, behind two much more computationally expensive methods.

	NN	FT	ST	DCG
Original Meshes	0.984	0.732	0.841	0.927
Classic MDS	0.969	0.731	0.833	0.922
Fast MDS	0.944	0.649	0.766	0.884
Least Squares MDS	0.992	0.863	0.938	0.969
Non-Metric MDS	0.991	0.853	0.929	0.965
GPS	0.749	0.453	0.582	0.745
Euclidean Random	0.975	0.770	0.868	0.936
Euclidean Normalised	0.978	0.793	0.884	0.943
Our method (join skeleton components)	0.986	0.843	0.932	0.963
Our method (delete smallest components)	0.986	0.844	0.933	0.964

in Table 2, and the precision-recall curves are shown in Figure 7. Our method achieves the third highest retrieval performance on this dataset, behind the least squares and non-metric MDS methods. This may be because the details matter less with this dataset, as the difference between the shape classes is much larger, and therefore keeping the details may only serve to increase the level of noise in the result. Using no canonical forms at all achieved a higher retrieval performance than the classic MDS, fast MDS and GPS methods, and achieved a higher nearest neighbour performance than the Euclidean distance based methods. Such a high performance from a rigid retrieval method shows that the different shape classes are easily distinguishable even when the non-rigid nature of the shapes is ignored. The precision-recall curves show that there is a noticeable gap in performance between our method, along with the least squares and non-metric MDS methods, and the others.

Table 2 shows the performance of both our methods for dealing with disconnected components. We achieve a tiny performance increase of 0.001 when only keeping the largest mesh component. This probably means that this is a better solution, but as there is only a small proportion of meshes with this problem (15 out of 1,200), they only make a minor impact on the overall performance.

The timings for each canonical form method on the non-rigid benchmark are shown in Table 3. The methods were run on a Linux PC with an Intel i7-3930K 3.2 GHz processor and 32 GB of memory. All methods are primarily implemented in Matlab, but with some parts in C++ for speed. The times for the four MDS based methods are for meshes which were simplified to approximately 2,000

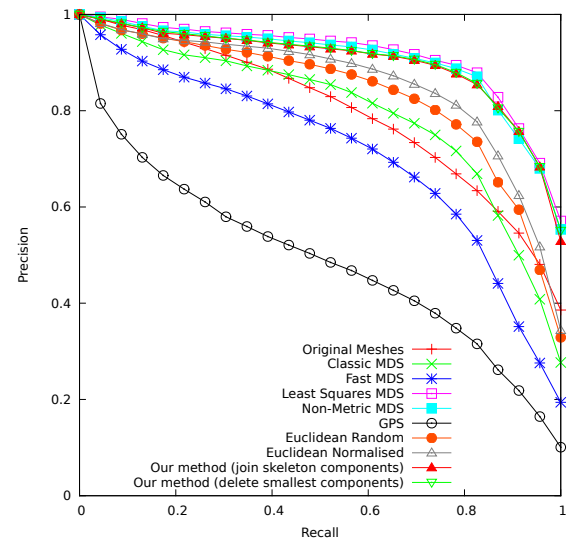


Fig. 7 Precision-recall curves for each method tested on the SHREC'15 non-rigid benchmark [18] using a view-based retrieval method [16]. Our method achieves the third best performance.

Tab. 3 Run-time of each method on the SHREC'15 non-rigid benchmark [18]. Our method achieves the second fastest run-time on this dataset. * These methods use meshes which have been simplified to $\sim 2,000$ vertices.

	Single mesh	Dataset
Classic MDS*	45 seconds	14 hours, 50 minutes
Fast MDS*	44 seconds	14 hours, 37 minutes
Least Squares MDS*	66 seconds	21 hours, 53 minutes
Non-Metric MDS*	104 seconds	34 hours, 44 minutes
GPS	2 seconds	44 minutes
Euclidean Random	23 seconds	7 hours, 37 minutes
Euclidean Normalised	23 seconds	7 hours, 39 minutes
Our method	11 seconds	3 hours, 36 minutes

vertices. This is because for a single full resolution mesh these methods take in excess of 20 minutes to compute the canonical form, and therefore the meshes must be simplified for these methods to finish within a reasonable length of time. Even with much lower resolution meshes, the MDS based methods are the slowest due to the use of geodesic distances. Our method is the second fastest of all the methods, being beaten in run-time by the GPS method. The GPS method however performed worst on all our retrieval experiments.

Our method therefore is significantly faster than the methods which achieved a slightly better retrieval performance, and achieves a significantly higher retrieval performance than the only method which had a faster run-time. We therefore achieve a very good trade-off between retrieval accuracy and efficiency.

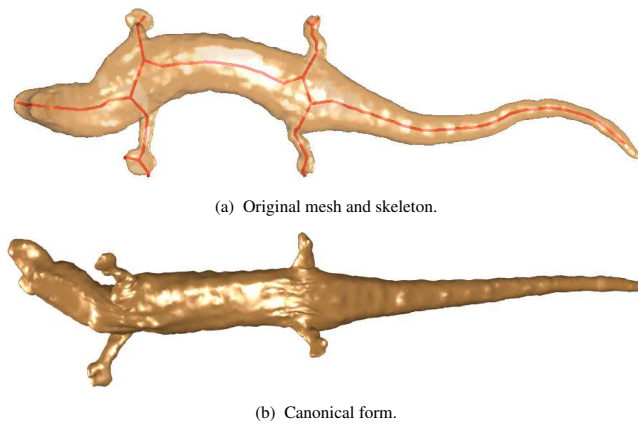


Fig. 8 Example of junctions which have not been correctly joined, and the impact on the canonical form.

4.3 Limitations

The skeleton refinement step described in Section 3.2.1 does not always merge junctions which are undesirably separate. An example of this is shown in Figure 8, where each of the arms and legs of the alligator are connected to the spine at a different junction. This is likely caused by the curved pose of the alligator's spine. This leads to the neck of the alligator not fully straightening out correctly, but even with this local inaccuracy we still achieve a retrieval nearest neighbour score of 1 for this mesh. There is room for future improvement to the skeleton refinement process, but this is a challenging problem as looser conditions for junction merging can lead to incorrect merging of junctions which should be separate.

The SHREC'15 non-rigid benchmark [18] contains some meshes with topological errors. We have already discussed meshes which contain multiple components in Section 4.2, but this dataset also contains meshes where parts of the mesh are undesirably joined together. Figure 9 shows an example mesh with this kind of topological error, and the resulting canonical form produced by each canonical form method we have tested. It can be seen that the arms of the manikin are incorrectly fused together along the forearms, which means that the arms are not correctly separated out by any of the canonical form methods. There are currently no canonical form methods that we are aware of which claim to be able to be able to correct for this level of topological error. The method by Boscaini *et al.* [5] proposes a method to handle errors where the incorrect connections are much smaller.

Our method is designed to work on objects which have a natural skeletal structure. Figure 10 shows a mesh from the “paper” class of the SHREC'15 non-rigid dataset. This mesh does not have a natural skeletal structure, and therefore our method fails to produce a sensible result. Our method

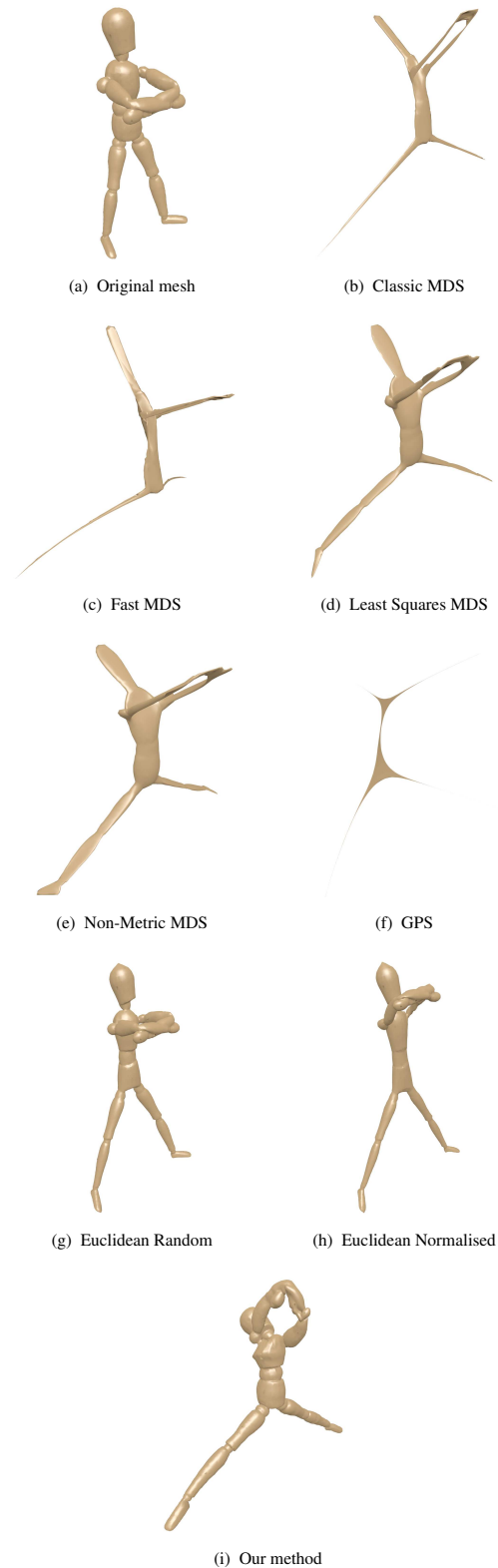


Fig. 9 Canonical forms for a mesh with incorrect connections.

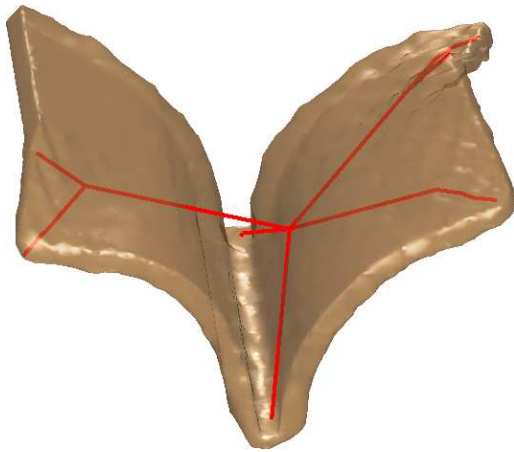


Fig. 10 Example of a mesh which does not have a natural skeleton structure.

would work for other man-made objects, as long as they have an obvious skeletal structure.

5 Conclusions

We have presented a novel method for computing the canonical form of a 3D mesh, which uses the mesh's skeleton to normalise its pose. We have shown that our method is able to achieve the same bending invariant pose as the previous state-of-the-art, whilst causing far less shape distortions than other methods. Our method is not able to correct for topological errors present in a mesh, and therefore there is room for future research in this direction. The retrieval performance produced using our canonical forms are competitive with other canonical form methods, achieving top performance on a recent canonical forms benchmark. Our method achieves high quality canonical forms, whilst achieving a significantly faster computation time over the previous state-of-the-art.

Acknowledgements

This work was supported by EPSRC Research Grant EP/J02211X/1.

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

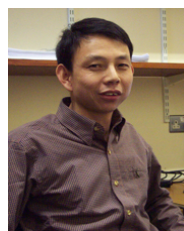
References

- [1] O. K.-C. Au, C.-L. Tai, H.-K. Chu, D. Cohen-Or, and T.-Y. Lee. Skeleton extraction by mesh contraction. In *ACM SIGGRAPH 2008 Papers*, SIGGRAPH, pages 44:1–44:10, New York, NY, USA, 2008. ACM.
- [2] R. A. Baeza-Yates and B. A. Ribeiro-Neto. *Modern Information Retrieval - the concepts and technology behind search, Second edition*. Pearson Education Ltd., Harlow, England, 2011.
- [3] M. Ben-Chen and C. Gotsman. Characterizing Shape Using Conformal Factors. In *Eurographics Workshop on 3D Object Retrieval*. The Eurographics Association, 2008.
- [4] I. Borg and P. Groenen. *Modern Multidimensional Scaling: Theory and Applications*. Springer Series in Statistics. Springer, 2005.
- [5] D. Boscaini, R. Girdziusas, and M. M. Bronstein. Coulomb Shapes: Using Electrostatic Forces for Deformation-invariant Shape Representation. In *Eurographics Workshop on 3D Object Retrieval*. The Eurographics Association, 2014.
- [6] E. Boyer, A. M. Bronstein, M. M. Bronstein, B. Bustos, T. Darom, R. Horaud, I. Hotz, Y. Keller, J. Keustermans, A. Kovnatsky, R. Litmany, J. Reininghaus, I. Sipiran, D. Smeets, P. Suetens, D. Vandermeulen, A. Zaharescu, and V. Zobel. SHREC '11: Robust Feature Detection and Description Benchmark. In *Eurographics Workshop on 3D Object Retrieval*. The Eurographics Association, 2011.
- [7] D.-Y. Chen, X.-P. Tian, Y.-T. Shen, and M. Ouhyoung. On visual similarity based 3D model retrieval. *Computer Graphics Forum*, 22(3):223–232, 2003.
- [8] K. Crane, C. Weischedel, and M. Wardetzky. Geodesics in heat: A new approach to computing distance based on heat flow. *ACM Trans. Graph.*, 32(5):152:1–152:11, Oct. 2013.
- [9] A. Elad and R. Kimmel. On bending invariant signatures for surfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(10):1285–1295, 2003.
- [10] A. Giachetti and C. Lovato. Radial symmetry detection and shape characterization with the multiscale area projection transform. *Computer Graphics Forum*, 31(5):1669–1678, 2012.
- [11] M. Hilaga, Y. Shinagawa, T. Kohmura, and T. L. Kunii. Topology matching for fully automatic similarity estimation of 3D shapes. In *Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH, pages 203–212, New York, NY, USA, 2001. ACM.
- [12] A. Johnson and M. Hebert. Using spin images for efficient object recognition in cluttered 3D scenes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(5):433–449, May 1999.
- [13] R. Kimmel and J. A. Sethian. Computing geodesic paths on manifolds. *Proceedings of the National Academy of Sciences*, 95(15):8431–8435, 1998.
- [14] B. Li, A. Godil, M. Aono, X. Bai, T. Furuya, L. Li, R. Lopez-Sastre, H. Johan, R. Ohbuchi, C. Redondo-Cabrera, A. Tatsuma, T. Yanagimachi, and S. Zhang. SHREC'12 Track: Generic 3D Shape Retrieval. In *Eurographics Workshop on 3D Object Retrieval*. The Eurographics Association, 2012.
- [15] Z. Lian, A. Godil, B. Bustos, M. Daoudi, J. Hermans, S. Kawamura, Y. Kurita, G. Lavou, H. V. Nguyen, R. Ohbuchi, Y. Ohkita, Y. Ohishi, F. Porikli, M. Reuter, I. Sipiran, D. Smeets, P. Suetens, H. Tabia, and

- D. Vandermeulen. SHREC '11 Track: Shape Retrieval on Non-rigid 3D Watertight Meshes. In *Eurographics Workshop on 3D Object Retrieval*. The Eurographics Association, 2011.
- [16] Z. Lian, A. Godil, X. Sun, and J. Xiao. CM-BOF: visual similarity-based 3D shape retrieval using clock matching and bag-of-features. *Machine Vision and Applications*, 24(8):1685–1704, 2013.
- [17] Z. Lian, A. Godil, and J. Xiao. Feature-preserved 3D canonical form. *International Journal of Computer Vision*, 102(1-3):221–238, 2013.
- [18] Z. Lian, J. Zhang, S. Choi, H. ElNaghy, J. El-Sana, T. Furuya, A. Giachetti, R. A. Guler, L. Lai, C. Li, H. Li, F. A. Limberger, R. Martin, R. U. Nakanishi, A. P. Neto, L. G. Nonato, R. Ohbuchi, K. Pevzner, D. Pickup, P. Rosin, A. Sharf, L. Sun, X. Sun, S. Tari, G. Unal, and R. C. Wilson. Non-rigid 3D Shape Retrieval. In *Eurographics Workshop on 3D Object Retrieval*. The Eurographics Association, 2015.
- [19] D. Pickup, X. Sun, P. L. Rosin, and R. R. Martin. Euclidean-distance-based canonical forms for non-rigid 3D shape retrieval. *Pattern Recognition*, 48(8):2500–2512, 2015.
- [20] D. Pickup, X. Sun, P. L. Rosin, R. R. Martin, Z. Cheng, Z. Lian, M. Aono, A. B. Hamza, A. Bronstein, M. Bronstein, S. Bu, U. Castellani, S. Cheng, V. Garro, A. Giachetti, A. Godil, J. Han, H. Johan, L. Lai, B. Li, C. Li, H. Li, R. Litman, X. Liu, Z. Liu, Y. Lu, A. Tatsuma, and J. Ye. Shape Retrieval of Non-Rigid 3D Human Models. In *Eurographics Workshop on 3D Object Retrieval*. The Eurographics Association, 2014.
- [21] D. Pickup, X. Sun, P. L. Rosin, R. R. Martin, Z. Cheng, S. Nie, and L. Jin. Canonical Forms for Non-Rigid 3D Shape Retrieval. In *Eurographics Workshop on 3D Object Retrieval*. The Eurographics Association, 2015.
- [22] M. Reuter, F.-E. Wolter, and N. Peinecke. Laplace-Beltrami spectra as Shape-DNA of surfaces and solids. *Computer-Aided Design*, 38(4):342–366, 2006. Symposium on Solid and Physical Modeling 2005.
- [23] K. Sfikas, T. Theoharis, and I. Pratikakis. Non-rigid 3D object retrieval using topological information guided by conformal factors. *The Visual Computer*, 28(9):943–955, 2012.
- [24] G. Shamai, M. Zibulevsky, and R. Kimmel. Accelerating the Computation of Canonical Forms for 3D Nonrigid Objects using Multidimensional Scaling. In *Eurographics Workshop on 3D Object Retrieval*. The Eurographics Association, 2015.
- [25] D. Smeets, J. Hermans, D. Vandermeulen, and P. Suetens. Isometric deformation invariant 3D shape recognition. *Pattern Recognition*, 45(7):2817–2831, 2012.
- [26] D. Smeets, J. Keustermans, D. Vandermeulen, and P. Suetens. meshSIFT: Local surface features for 3D face recognition under expression variations and partial data. *Computer Vision and Image Understanding*, 117(2):158–169, 2013.
- [27] J. Sun, M. Ovsjanikov, and L. Guibas. A concise and provably informative multi-scale signature based on heat diffusion. *Computer Graphics Forum*, 28(5):1383–1392, 2009.
- [28] J. W. Tangelder and R. C. Veltkamp. A survey of content based 3D shape retrieval methods. *Multimedia Tools and Applications*, 39(3):441–471, 2008.
- [29] X.-L. Wang and H. Zha. Contour canonical form: An efficient intrinsic embedding approach to matching non-rigid 3D objects. In *Proceedings of the 2nd ACM International Conference on Multimedia Retrieval*, ICMR '12, pages 31:1–31:8, New York, NY, USA, 2012. ACM.
- [30] H.-B. Yan, S.-M. Hu, R. Martin, and Y.-L. Yang. Shape deformation using a skeleton to drive simplex transformations. *IEEE Transactions on Visualization and Computer Graphics*, 14(3):693–706, May 2008.
- [31] X. Ying, S.-Q. Xin, and Y. He. Parallel Chen-han (PCH) algorithm for discrete geodesics. *ACM Trans. Graph.*, 33(1):9:1–9:11, Feb. 2014.



David Pickup obtained an MEng degree from the University of Bristol and a PhD from the University of Bath. He is currently a post-doc research associate within the Visual Computing research group at Cardiff University. His current research centres on three-dimensional shape retrieval.



Xianfang Sun received his PhD from the Institute of Automation, Chinese Academy of Sciences in 1994. He is currently a lecturer at the School of Computer Science, Cardiff University, Wales, UK. His main research interests include computer vision, computer graphics, pattern recognition and artificial intelligence.



Paul L. Rosin is Professor at the School of Computer Science & Informatics, Cardiff University. Previous posts include lecturer at the Department of Information Systems and Computing, Brunel University London, UK, research scientist at the Institute for Remote Sensing Applications, Joint Research Centre, Ispra, Italy, and lecturer at Curtin University of Technology, Perth, Australia.

His research interests include the representation, segmentation, and grouping of curves, knowledge-based vision systems, early image representations, low level image processing, machine vision approaches to remote sensing, methods for evaluation of approximation algorithms, etc., medical and biological image analysis, mesh processing, non-photorealistic rendering and the analysis of shape in art and architecture.



Ralph R. Martin obtained his PhD in 1983 from Cambridge University. Since then he has been at Cardiff University, where he now holds a Chair and leads the Visual Computing research group. He is also a Guest Professor at Tsinghua and two other Universities in China, and is Director of Scientific Programmes of the One Wales

Research Institute of Visual Computing. His publications include over 250 papers and 14 books covering such topics as solid modelling, surface modelling, reverse engineering, intelligent sketch input, mesh processing, video processing, computer graphics, vision based geometric inspection, and geometric

reasoning. He is a Fellow of the Learned Society of Wales, the Institute of Mathematics and its Applications, and the British Computer Society. He is on the editorial boards of "Computer Aided Design", "Computer Aided Geometric Design", "Geometric Models", and other journals.